Uncountability of the reals (includes a lot of bs that needs to be fixed)

October 5, 2014

A set is said to be countable (or countably infinite) if there is a bijection (one-to-one correspondance) between its elements and the natural numbers. For example, the integers $\mathbb{Z}$ are countable because we can assign the following mapping from $\mathbb{N}$:

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{-n+1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

A set is said to be uncountable if its cardinality is greater than that of the natural numbers. Even though there is an infinite number of natural numbers, it's possible to show that, for example, the set of real numbers is strictly larger than the set of natural numbers. One such proof is Cantor's diagonalisation argument: list the real numbers between 0 and 1 in some order. Then form a number $y \in [0, 1)$ in the following way. For each number x_k, take the kth digit after the decimal point. If it's a 5, then the kth digit of y is a 4, otherwise the kth digit of y is a 5. You can see that y cannot appear in the list of numbers previously given. Clearly y is a real number between 0 and 1. Therefore the set of real numbers has a higher cardinality than the natural numbers.

Now, the real numbers can be partitioned in many ways. For example, they are the union of the negative reals and the nonnegative reals. Another way to put it is that it is the union of the rational numbers and the irrational numbers. We can show that the rational numbers are countable, again by giving a bijection to the natural numbers (list the rational numbers in a table where the column represents the numerator and the row represents the denominator, and traverse it starting from the top left, diagonally going toward the bottom left, one diagonal at a time; some numbers will be listed twice, e.g. 1/2 and 2/4, but this still shows the set being countable). Thus the reals are uncountable strictly because of the irrational numbers, and thus the irrational numbers are uncountable.

We can also say that the reals are the union of the algebraic numbers and the transcendental numbers. The algebraic numbers are the numbers that are roots of polynomial equations with integer coefficients. For example, $x^2 - 2 = 0$ gives $x = \sqrt{2}$, so $\sqrt{2}$ is an algebraic number. The definition of the transcendental numbers is precisely the reals that are not algebraic. Examples of transcendental numbers are π, e, and Champernowne's constant. It can be shown that the algebraic numbers are countable: think of the coefficients of polynomial equations as ordered tuples, and the proof is similar to that of the countability of the rationals (which are pairs: tuples of size 2). Then the reals are uncountable strictly because of the transcendental numbers.

You've no doubt heard of people who memorized hundreds of digits of pi, or of computers that found the trillionth digit of pi and whatnot. The reason this is possible is that pi is *computable*. There are ways to determine pi to any precision we desire. For example, if we generate random points uniformly in a square where $x \in [0, 1]$ and $y \in [0, 1]$ and then count the proportion of points that satisfy the equation $\sqrt{x^2 + y^2} < 1$, then we get an approximation for $\frac{\pi}{4}$. This is an example of a Monte Carlo method. Another approximation for $\frac{\pi}{4}$ given by Leibniz is the sum $\sum_{i=0}^{n} \frac{(-1)^i}{2i+1}$.

Now, it can be shown that there are countably many algorithms. The proof is the same as that of the countability of the algebraic numbers: treat each algorithm as a tuple. Since to compute a number like π or e to some arbitrary precision we use an algorithm, it follows that there is a countable number of computable numbers. Hence the reals are uncountable strictly because of uncomputable numbers.

Thus we have a new partition of the reals: the computable numbers and the uncomputable numbers. This is getting hairy. What's an uncomputable number? One of the most important computer science problems in the 20th century was the halting problem. It was proven by Alan Turing that no algorithm could be given that would take as input another algorithm, and output whether this algorithm halts on any given input. That is, the halting problem is undecidable. Gregory Chaitin devised a number, Ω, called Chaitin's constant, that represents the probability that a random algorithm will halt, given an input. Because the halting problem is undecidable, Chaitin's constant is uncomputable. Then the reals are uncountable (yes, even after all this, the diagonalization argument of Cantor still holds) strictly because of uncomputable numbers. You can see that most of the numbers that we deal with in applications are in a countable set, and there is even a field, called computable analysis, that is concerned with just such things.

Are we done? Well, not really. Notice that the above paragraph is finite in length (otherwise you wouldn't have made it to this paragraph without skipping it, in which case you couldn't comment on its length since you didn't read it). Clearly the definition that we gave for Chaitin's constant is finite. Then we can make a similar argument that was made for algorithms, and apply it to definitions of numbers. This proves that there is a countable number of definable numbers, like Chaitin's constant, so uncomputables are not the culprits that we are looking for. Indeed, the real numbers are uncountable strictly because of indefinable numbers. We *literally can't even* talk about those.

What's more? The kind of things taught in introductory analysis - sequences and series - only make sense over countable indices. Here's an example given by Terrence Tao at the start of his book on measure theory. Suppose that we have some function ϕ such that $\phi(a) \geq 0$ for all a, and a is taken from some possibly uncountable set A. Take the series

$$\sum_{a \in A} \phi(a).$$

Suppose that it is finite. That is, this series equals a real number. Then there can be at most a countable number of nonzero elements in A: since the series is finite, for any n there is at most a finite number of elements of $a \in A$ such that $\phi(a) \geq \frac{1}{n}$, because $\frac{1}{n} \cdot \infty = \infty$. So the set $\left\{a \in A : \phi(a) > \frac{1}{n}\right\}$ is finite for any n. Let $A' = \{a \in A : \phi(a) > 0\}$. Then

$$A' = \bigcup_{n=1}^{\infty} \left\{a \in A : \phi(a) > \frac{1}{n}\right\}.$$

A countable union of finite sets is countable, and the result follows.

A popular example in probability for something that is counterintuitive is related to this. Suppose that you have a dart board. What's the probability that a dart lands on any specific point? It must be zero! Otherwise, suppose that it's some small number ϵ. There are infinitely many points on the board, so the probability of hitting the board anywhere is $\epsilon \cdot \infty = \infty$. This makes no sense, since probabilities are between 0 and 1. So when dealing with continuous domains (the real numbers), it only makes sense to ask things like what's the probability that a dart will land in the upper half of the board.